Economics, Law and Policy
ISSN 2576-2060 (Print) ISSN 2576-2052 (Online)
Vol. 4, No. 1, 2021
www.scholink.org/ojs/index.php/elp

Original Paper

Near Identification of Policy Trade-Offs Using Preference Proxies to Elicit Hidden Information in the Linear Case

Jean-Paul Azam[1]

[1] Toulouse School of Economics, Toulouse, France

Received: May 1, 2021 Accepted: May 16, 2021 Online Published: June 1, 2021
doi:10.22158/elp.v4n1p48 URL: http://dx.doi.org/10.22158/elp.v4n1p48

Abstract

This paper shows that neither OLS nor 2SLS can generically identify policy trade offs in the linear case, except under extreme assumptions. Practitioners must be content with near identification and the paper discusses how to choose between these two methods. It shows that a two-stage approach using preference proxies to elicit hidden information can potentially narrow the identification gap and that a simple specification test can be used to assess whether these proxies really contribute to improving identification.

Keywords

policy tradeoffs, near identification, preference proxies

1. Introduction

This paper provides a simple framework for discussing the choice between OLS and two-stage (2SLS/Control Function) approaches for estimating and testing policy effectiveness using non-experimental data in a linear model. From a practical point of view this choice is not innocuous, as these two approaches often lead to opposite conclusions that can influence the welfare of millions of people. A good example of such a dilemma is provided by the so-called “aid-ineffectiveness” literature. Boone (1996) found that foreign aid does not affect economic growth in recipient countries, using OLS. This diagnosis became part of the conventional wisdom about this issue as illustrated by popular books like Easterly (2006) and Moyo (2009). In contrast, Arndt et al. (2015) revisited the issue and found instead that foreign aid is in fact effective at boosting economic growth and other desirable outcomes in recipient countries, using Instrumental Variables (IVs). However, heated debates are taking place among practitioners regarding the use of IVs, as illustrated by a popular blog that hosted a contribution with the title: “Friends don’t let Friends do IV” (Angus, 2015). The present paper tries to cool the debate down within a near-identification framework.

Unlike randomized control trials, real-world data require the joint modeling of the policy trade off under study and the policy maker's decisions. This entails generically that the policy trade off cannot be perfectly identified using least squares (OLS or 2SLS/Control function), as shown below, except under special circumstances. Because policy decisions will be made anyway, econometricians must take a more pragmatic approach to inform policy makers and voters using whatever method performs best. The scope of the discussion below is limited to the choice between these two most common estimation methods. It uses a near-identification framework, first discussed by Fisher (1965). This approach is based on the intuition that "information on the variances of the disturbance terms of a multiple equations system can be used for identification of the equation with the smallest disturbance variance" (Fisher, 1965, p. 409). However, policy effectiveness analysis yields only qualitative information about this, and empirical testing is required to assess the validity of the implicit ranking of disturbance variances across equations. Moreover, the OLS and 2SLS/Control function estimation methods generate different disturbance terms, and hence yield different outcomes in terms of near identification. The present paper shows how these differences may be used for choosing among these two methods the one that comes closer to proper identification.

The next section presents the linear model used and the identification problem faced when using OLS. Section 3 shows that a two-stage approach using proxies aimed at capturing the policy maker's preference parameters may improve the analysis and suggests a test that can be used to evaluate if these preference proxies have really contributed to narrow the identification gap or not. Section 4 briefly concludes.

2. The Model and its Identification Problem

2.1 The Setting

The econometrician wants to evaluate "policy effectiveness" in a given domain. The impact of a policy *p*, which is here a continuous variable, on an outcome variable of interest *y*, is embedded in the following linear equation in which *x* is an exogenous shift (control) variable (or a list of several of them) aimed at providing a *ceteris paribus* policy effectiveness diagnosis. The econometrician does not observe variables *e* and ε that also affect the outcome:

$$y = \alpha + \beta x + \gamma p + \delta e + \varepsilon . \tag{1}$$

Unless otherwise specified, all the Greek-letter parameters are defined as positive. For the sake of simplicity, assume that $E(e) = E(\varepsilon) = 0$; if they had non-zero expected values, the latter would simply be added to the intercept to yield the equivalent specification with a different α. The policy effectiveness claim is that γ is nonzero and works in the desirable direction, say $\gamma > 0$ for the sake of concreteness. To test it, the econometrician will seek to identify (1) as closely as possible to put himself in a position to interpret his empirical findings.

It would be unscientific to assume without further testing that the policymaker is incompetent and stupid. The "Public Choice revolution" of the 1960s, initiated by people like James Buchanan and Gordon Tullock, has exerted a sufficient influence on the profession to preclude the credible use of this type of assumptions without testing. Let us make instead the following three assumptions:

(i) ***Asymmetric Information*****:** The policy maker observes x, p and e before making her decision**,** and then she observes y ex post, while the econometrician only observes y, x and p ex post.

(ii) ***Efficient Information Processing*****:** The policy maker makes the best use of her information so that $E(e\,\varepsilon) = 0$.

(iii) ***Quasi-Concave Preferences*****:** The policy maker seeks to maximize the following objective function:

$$\max_p R \equiv (\theta + E(y))^2 - p^2 \text{ s.t. } (1), \tag{2}$$

where $E(y)$ is a shorthand notation for $E(y \mid x, p, e)$. Assumptions (i) and (ii) are natural to make for any rational-choice scientist. Assumption (iii) is not very general, but it is chosen because it yields convenient predictions that keep the resulting econometric specifications linear and tractable. In this specification, θ is a preference parameter of the policy maker, which is her private information. A higher θ means that she is more sensitive to the outcome variable y, as:

$$\frac{\partial^2 R}{\partial \theta \, \partial E(y)} = 2 . \tag{3}$$

This objective function is not concave, but given the linear constraint (1), a quasi-concave objective function is sufficient to determine the optimum, if the latter exists. Figure 1 depicts a case where the optimum is easily seen to exist, with $\alpha + \beta x + \delta e > 0, \theta > 0$ and $0 < \gamma < 1$. These conditions entail that $R > 0$ in this case. They are not necessary for the optimum to exist in general and other cases can easily be worked out.

In Figure 1, the upward-sloping line labeled $E(y)$ represents (1) with ε set to zero. The upward-sloping convex curve depicts an indifference curve derived from the objective function (2). The latter is quasi-concave because this indifference curve is convex in this case. This can be checked by computing the total differential of R and setting it equal to zero, and then, by rearranging the terms and taking the derivative with respect to p a second time to yield:

$$\frac{\partial E(y)}{\partial p} = \frac{p}{\theta + E(y)} > 0 \text{ and } \frac{\partial^2 E(y)}{\partial p^2} = \frac{R}{(\theta + E(y))^3} > 0 . \tag{4}$$

This entails that the set of points that are preferred to all the points of an indifference curve is convex. The optimum point is classically found where the indifference curve is tangent to the constraint. The first-order condition for maximizing (2) reads:

$$p^* = \gamma(\theta + E(y)^*) . \tag{5}$$

Hence, the policy-maker's optimum choice is found at the intersection of (1) with a straight line whose equation may be written as $E(y) = p/\gamma - \theta$. This line is represented in figure 1 as the steeper upward-sloping line labeled $p/\gamma - \theta$. It is now obvious that the policy-maker's optimum exists in the case of Figure 1. This diagram may be used to show that the two components of the policy maker's informational advantage over the econometrician, namely θ and e, are playing in opposite directions regarding the identification of $E(y)$.

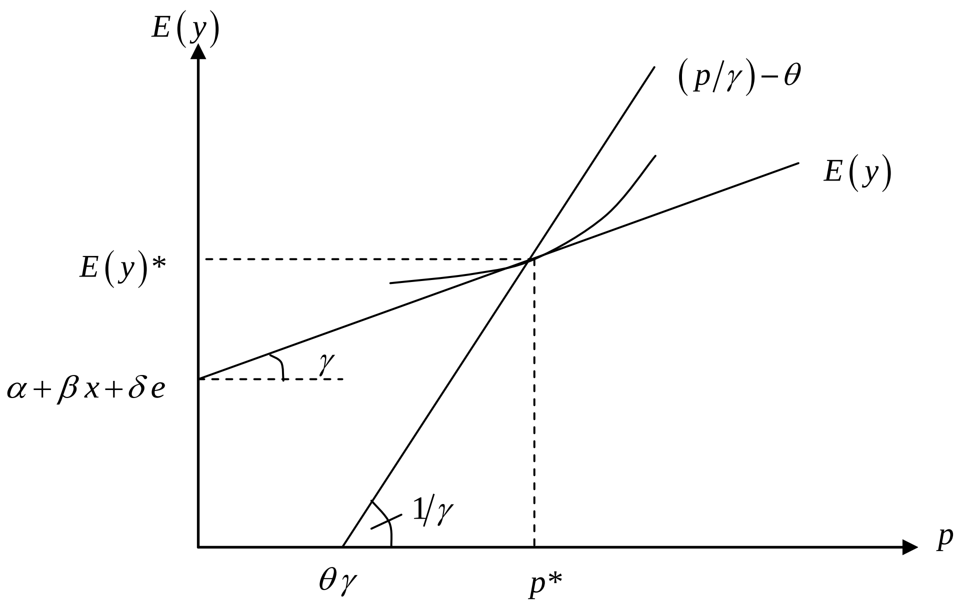

Figure 1. Policy-Maker's Optimum with $0 < \gamma < 1$

Figure 1 may be used to show that changes in the policy maker's preferences play a key part in identifying $E(y)$. The impact of a *ceteris paribus* increase in θ would entail a rightward shift of the $(p/\gamma) - \theta$ line. Hence, unobservable variations in θ are tracing out the $E(y)$ line. In contrast, the unobservable variations in e are easily shown to lie at the core of this identification problem. For example, an unobserved increase in *e*, given $\delta > 0$, would entails an upward shift of the $E(y)$ line. Hence, the variations in *e*, given the other variables and parameters, would in fact trace out the $(p/\gamma) - \theta$ line rather than the $E(y)$ one. Still, the impacts of these two unobservable variables are difficult to single out because they both entail an increase in both $p*$ and $E(y*)$. This similarity of impacts comes out clearly from deriving formally the policy rule governing $p*$.

Substituting for $E(y)*$ in the first-order condition (5) and rearranging the terms allows us to derive the reduced-form equation for $p*$ that describes the policy-maker's policy rule:

$$p^* = \frac{\gamma\,\alpha}{1-\gamma^2} + \frac{\gamma\,\beta}{1-\gamma^2}\,x + \frac{\gamma\,\theta}{1-\gamma^2} + \frac{\gamma\,\delta}{1-\gamma^2}\,e\,. \tag{6}$$

This policy rule includes both components of the policy maker's informational advantage over the econometrician, namely θ and e, while it does not include ε. Moreover, its coefficients are mongrel parameters of the relevant parameters of the structural equation (1) and the policy maker's objective function (2). Still, it turns out that (6) provides the basis of the solution to mitigate the identification problem explained in the next sub-section.

2.2 The Identification Problem

Now, assume that the econometrician has a sample of observations of y, x and p, as well as a few other variables used below, indexed implicitly by $i \in S$. Given the information available to him, the econometrician tries to estimate:

$$y = \alpha + \beta\,x + \gamma\,p + f\,, \tag{7}$$

where $f \equiv \delta\,e + \varepsilon$. This "random disturbance" term is in fact a function of p, x and θ because one can substitute for $\delta\,e$ from (6) after rearranging the terms to write:

$$\delta\,e = \left(\frac{1-\gamma^2}{\gamma}\right)p - \alpha - \theta - \beta\,x\,. \tag{8}$$

Therefore, substituting (8) into (6) yields a second relation between y and p (and ε as well) :

$$y = \left(1/\gamma\right)p - \theta + \varepsilon\,. \tag{9}$$

Notice that the difference between the coefficients of p in (9) and in (7) is:

$$\frac{1}{\gamma} - \gamma = \frac{1-\gamma^2}{\gamma}\,. \tag{10}$$

This measures the maximum potential bias that would result from estimating (9) rather than (7), neglecting the omitted-variable biases that would be caused by the absence of e in estimating (7) and of θ in estimating (9).

If there are several policymakers involved in producing the outcomes captured by this sample, or if the policy maker's preferences vary over time, then unobserved θ will in general vary across $i \in S$. Let us define $\overline{\theta}$ as its mean and $\eta \equiv \theta - \overline{\theta}$ as its deviations from the mean, with $\sum_{i\in S}\eta = 0$. Then (9) may be written as:

$$y = -\overline{\theta} + \left(1/\gamma\right)p + \left(\varepsilon - \eta\right). \tag{11}$$

There is no reason to expect that these deviations from the mean of the policy makers' preference parameters will be correlated with ε and it is thus natural to assume that $E\left(\varepsilon\,\eta\right) = 0$. Then the following identification failure proposition can be proved simply.

***Proposition* 1:** Equation (7) cannot in general be identified by OLS as the latter will select a linear combination of (7) and (11) instead with a non-zero weight attached to either one.

***Proof*:** The idea of the proof is to show that a linear combination of (7) and (11) will in general have a smaller sum of squared residuals than those resulting from estimating either one separately, if the sample is large enough. Define $\lambda \in [0,1]$. Then we can write in obvious notation a linear combination of (7) and (11) which has the same structure as (7):

$$\begin{aligned} y &= \lambda\, y_{(7)} + (1-\lambda)\, y_{(11)} = \\ &\left(\lambda\,\alpha - (1-\lambda)\,\overline{\theta}\right) + \lambda\,\beta\,x + \left(\lambda\,\gamma + (1-\lambda)/\gamma\right) p + \varepsilon - \eta + \lambda\left(\eta + \delta\,e\right) \end{aligned} . \quad (12)$$

It follows from this identical structure that identification cannot result from structural parameters restrictions and must be assessed by looking at the residual variances (Note 1). Then, OLS will choose λ (and other parameters) such that:

$$\begin{aligned} \min_{\lambda} \sum\nolimits_{i\in S} \left(\varepsilon - \eta + \lambda(\eta + \delta\,e)\right)^2 &= \sum\nolimits_{i\in S}\left(\varepsilon^2 + \eta^2\right) + \lambda^2 \sum\nolimits_{i\in S}\left(\eta^2 + \delta^2 e^2 + 2\,\delta\,\eta\,e\right) \\ &+ 2\,\lambda \sum\nolimits_{i\in S}\left(\left(\varepsilon\,\eta + \delta\,e(\varepsilon-\eta) - \eta^2\right)\right) - 2\sum\nolimits_{i\in S}\varepsilon\,\eta \end{aligned} . \quad (13)$$

Taking due account that $E(\varepsilon\,\eta) = E(\varepsilon\,e) = E(\eta\,e) = 0$, the expected value of the sum to be minimized is thus $E\left(\varepsilon^2\right) + E\left(\eta^2\right) + \lambda^2\left(E\left(\eta^2\right) + \delta^2 E\left(e^2\right)\right) - 2\lambda\,E\left(\eta^2\right)$. Then, given a large enough sample, OLS will select:

$$\lambda_{OLS} = \frac{E\left(\eta^2\right)}{E\left(\eta^2\right) + \delta^2 E\left(e^2\right)} . \quad (14)$$

Hence, assuming realistically that $E\left(\eta^2\right) \ll \infty$, OLS will only identify perfectly the policy trade off if $\delta^2 E\left(e^2\right) = 0$, i.e., if either the policy maker has no information advantage over the econometrician about e, or for some reason decides not to use it. This cannot be assumed without testing.

Proposition 1 simply provides the econometric equivalent of the geometrical result of the previous sub-section: OLS will choose to estimate a linear combination of (7) and (11) to minimize the impact of a weighted sum of unobserved variations of e, θ and ε on the residuals. In other words, our identification failure diagnosis is mainly driven by an omitted variable bias, the omitted variable being in this case a piece of information used by the policy maker but unobservable to the econometrician whose variance is $E\left(e^2\right)$. This comes out clearly from defining the OLS identification gap as:

$$1 - \lambda_{OLS} = \frac{\delta^2 E\left(e^2\right)}{E\left(\eta^2\right) + \delta^2 E\left(e^2\right)} . \quad (15)$$

This gap will only fall to zero if $\left(\delta^2 E\left(e^2\right) / E\left(\eta^2\right)\right) \to 0$, i.e., if the policy maker has no information advantage about e over the econometrician, an unrealistic assumption. The identification gap will not be a big deal if the unobserved variations are dominated by those of the policy maker's preferences; it will be large if they are instead dominated by the latter's information advantage about e over the econometrician. This is the essence of the problem raised by the endogeneity of the policy variable.

Having understood that this identification problem is due to an omitted variable, we can now look for a good proxy that could be substituted to unobserved *e* to mitigate this problem and provide a fairly good estimate of (1). The main lead to find it is obvious from either Figure 1 or a glance at (6). We observe that $p*$ is responding to increases in e or θ by a seemingly unexplained increase, given the variables observed by the econometrician. Then, the challenge is to disentangle the impacts of these two unobserved variables to elicit the impacts of changes in e after controlling somehow for the changes in θ. The next section shows how preference proxies can be used to make some progress in this direction, and how this progress can be assessed statistically.

3. Using Preference Proxies to Elicit Hidden Information

The way to use the information provided by the deviations of the policy maker's behavior relative to the latter's fitted value using regressors equally known to the econometrician and the policy maker to improve identification is a direct extension of Hausman (1978) and Nakamura and Nakamura (1981) in the simplest case. It has since been named the control-function approach, massively generalized, and applied in a very large number of papers. This section customizes its application in its simplest form to the problem at hand in two steps.

3.1 Signal Extraction and Estimation

Assume now that the econometrician's data set includes one or more variables w that are liable to be jointly correlated with θ, and thus labeled preference proxies, while they are not included in *x*. Let our econometrician assume that:

$$\theta = \pi + \mu w + \zeta, \text{ with } E(\zeta) = E(\zeta\,\varepsilon) = 0. \tag{16}$$

Notice that the signs of $\{\pi, \mu\}$ are unknown. As θ is not observable directly, (16) cannot be directly tested. However, it can be tested indirectly as shown below. The econometrician can use w as a preference proxy (PP) for θ in the policy rule. The first-stage equation is obtained by substituting for θ from (16) into (6) to read:

$$p = \hat{a} + \hat{b}\,x + \hat{c}\,w + \hat{g} \tag{17}$$

Table 1 shows what each of these estimated coefficients of (17) is estimating in the present framework. These are complicated mongrel parameters whose exact significance and statistical properties are far from obvious because the expected value of a ratio is not equal to the ratio of the expected values of its numerator and its denominator. Moreover, the correct standard errors of the underlying parameters are anybody's guess. It seems difficult to assume that anything useful could be tested from these parameters regarding identification of the policy trade off. It will become clearer below that the key requirement is that $\hat{g}$ be orthogonal to x and w, i.e., that β and μ be nonzero, and not the statistical significance of their parameters in (17). In particular, the maximum potential bias

$\left(1-\gamma^2\right)/\gamma$ appears in the denominator of all the coefficients, which might thus look less significant, the more needed is this two-stage approach to reduce the identification gap. Unfortunately, however, performing some tests on this type of parameters is common practice in the profession as a check on the weakness of instrumental variables (see, e.g., Stock & Yogo, 2005).

Table 1. The First-Stage Mongrel Parameters

$\hat{a}$	$\hat{b}$	$\hat{c}$	$\hat{g}$
$\frac{\gamma(\alpha+\pi)}{1-\gamma^2}$	$\frac{\gamma\beta}{1-\gamma^2}$	$\frac{\gamma\mu}{1-\gamma^2}$	$\frac{\gamma}{1-\gamma^2}(\delta e+\zeta)$

Note. This table shows the correspondence between the parameters of (17) and the deeper parameters of the model.

Then, our econometrician now includes $\hat{g}$ in his second-stage equation as in:

$$y=\alpha+\beta x+\gamma p+\phi\hat{g}+\nu . \tag{18}$$

Proposition 2 follows.

***Proposition* 2:** OLS applied to (15) will estimate a linear combination of (1) and (11), giving the former a weight equal to:

$$\lambda_{PP}=\frac{E\left(\eta^2\right)+E\left(\varepsilon^2\right)}{E\left(\eta^2\right)+2E\left(\varepsilon^2\right)}>\frac{1}{2}. \tag{19}$$

***Proof*:** Notice first that ν will be an estimate of ε because the inclusion of $\hat{g}$ is controlling for $\delta e+\zeta$, so that $\sum_{i\in S}\nu^2$ will be an estimate of $\sum_{i\in S}\varepsilon^2$. Then, using the same approach as in the proof of proposition 1, as well as the identifying parameter restriction that w is excluded from (11), we can write the linear combination of (18) and (11) as:

$$\begin{aligned} y=\lambda y_{(18)}+(1-\lambda)y_{(11)}&=\left(\lambda\alpha-(1-\lambda)\bar{\theta}\right)\\ &+\lambda\beta x+\left(\lambda\gamma+(1-\lambda)/\gamma\right)p+\lambda\phi g+\lambda\nu+(1-\lambda)(\varepsilon-\eta) \end{aligned}. \tag{20}$$

Its structure is identical to that of (18), so that we again need to look at the random disturbance terms to assess identification. The sum of squared residuals of (20) reads:

$$\begin{aligned} \sum_{i\in S}\left(\lambda\nu+(1-\lambda)(\varepsilon-\eta)\right)^2&=\sum_{i\in S}\left(\lambda^2\nu^2+(1-\lambda)^2\left(\varepsilon^2+\eta^2-2\varepsilon\eta\right)\right)\\ &+2\sum_{i\in S}\lambda\,\nu(1-\lambda)(\varepsilon-\eta) \end{aligned}. \tag{21}$$

Then, because $E\left(\nu^2\right)=E\left(\varepsilon^2\right)$ and $E(\varepsilon\eta)=E(\varepsilon-\eta)=E(\nu)=0$, the expected value of the sum to be minimized is equal to $\lambda^2 E\left(\varepsilon^2\right)+\left(1-\lambda\right)^2\left(E\left(\varepsilon^2\right)+E\left(\eta^2\right)\right)$. The latter is minimal when (19) holds.

The resulting identification gap is:

$$1-\lambda_{PP}=\frac{E\left(\varepsilon^2\right)}{E\left(\eta^2\right)+2E\left(\varepsilon^2\right)}<\frac{1}{2}. \qquad (22)$$

Hence, this two-stage approach produces an identification gap that only falls to zero if $\left(E\left(\varepsilon^2\right)/E\left(\eta^2\right)\right)\rightarrow 0$, i.e., if the policy maker benefits from a perfect information or if there is an infinite variance of the policy maker's preference deviations from the mean. Both assumptions are unrealistic, so that this method also results in an identification gap. Then, the relevant question to ask is under which conditions the preference-proxies method results in a smaller identification gap than OLS.

3.2 Comparison of the Two Approaches

A glance at (15) and (22) shows that the OLS identification gap depends on the extent of the asymmetric information about e between the policy maker and the econometrician while the two-stage PP one depends on the symmetric uncertainty common to the two players. Moreover, we see that the two-stage PP approach always gives a larger weight to (1) than to (11), while OLS does not.

To go deeper into this comparison of the two approaches, assume that the econometrician wants to choose the estimation method that yields the lower identification gap. Then, Figure 2 shows how much lower than $E\left(\varepsilon^2\right)$ must $\delta^2 E\left(e^2\right)$ be for OLS to be preferred to PP, for a given value of $E\left(\eta^2\right)$. The *y*-axis represents the level of λ while the values of $E\left(\varepsilon^2\right)$ and $\delta^2 E\left(e^2\right)$ are measured on the *x*-axis. The diagram is read as follows: for any level of $E\left(\varepsilon^2\right)$, the upper curve shows the corresponding value of λ_{PP}. Given the latter, the lower curve shows the maximum value of $\delta^2 E\left(e^2\right)$ such that OLS yields a weakly lower identification gap. An example of this determination is shown by the arrow-bearing lines that intersect the two curves at the same λ level, i.e., for the same identification gap.

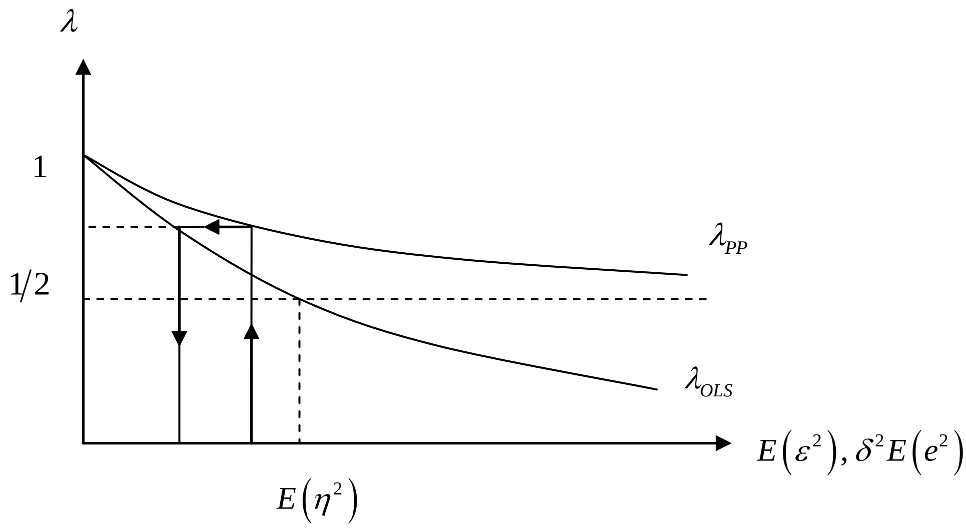

Figure 2. Determination of the Preferred Estimation Method

This intuition is captured more formally in proposition 3.

***Proposition* 3:** The preference proxies approach dominates OLS, i.e., $\lambda_{PP} > \lambda_{OLS}$, iff:

$$\delta^2 E\left(e^2\right) > \frac{E\left(\varepsilon^2\right) E\left(\eta^2\right)}{E\left(\varepsilon^2\right) + E\left(\eta^2\right)}. \tag{23}$$

***Proof*:** Inequality (23) simply results from subtracting (14) from (19) and requiring that the result be larger than zero, then re-arranging the terms.

***Comment*:** The right-hand side of (23) is smaller than $\min\left\{E\left(\varepsilon^2\right), E\left(\eta^2\right)\right\}$. Hence, given the policy maker's informational advantage about e, PP is more likely to be preferred to OLS, the better the goodness of fit of (18).

The next section shows how this theoretical finding can be made operational empirically. It shows that the Hausman test, as reformulated by Nakamura and Nakamura (1981), is providing a quantitative index that enables the econometrician to conclude empirically whether the preference proxies used have made a significant contribution to improve the near identification of the policy trade off or not.

3.3 A Test for Assessing Improved Identification

To simplify notation, let us define $\hat{\lambda} \equiv \lambda_{OLS}$ and $\tilde{\lambda} \equiv \lambda_{PP}$ and $\hat{f}$ and $\tilde{f}$ as the relevant residuals. Then, we may write (12) and (20) respectively as:

$$y = \left(\hat{\lambda}\alpha - \left(1-\hat{\lambda}\right)\overline{\theta}\right) + \hat{\lambda}\beta x + \left(\hat{\lambda}\gamma + \left(1-\hat{\lambda}\right)/\gamma\right)p + \hat{f} \tag{24}$$

and

$$y = \left(\tilde{\lambda}\alpha - \left(1-\tilde{\lambda}\right)\overline{\theta}\right) + \tilde{\lambda}\beta x + \left(\tilde{\lambda}\gamma + \left(1-\tilde{\lambda}\right)/\gamma\right)p + \tilde{\lambda}\phi\hat{g} + \tilde{f}\ . \tag{25}$$

For the sake of simplicity also, let us define $\hat{p}$ as the shorthand notation for the fitted part of the first-stage equation (17):

$$p = \hat{p} + \hat{g} \tag{26}$$

With this simplified notation it is straightforward to prove the following proposition:

***Proposition* 4:** Estimating (25) by OLS yields a natural estimator of the contribution of the preference proxies as the coefficient of $\hat{g}$ is an estimate of:

$$\tilde{\lambda}\phi = \left(\tilde{\lambda} - \hat{\lambda}\right)\left(\frac{1-\gamma^2}{\gamma}\right). \tag{27}$$

***Proof*:** The proof is a direct extension of Nakamura and Nakamura (1981). Definition (26) implies that $\hat{g}$ is orthogonal to $\hat{p}$ and hence to x. Now, substitute (26) for p in both (24) and (25) to yield respectively:

$$y = \left(\hat{\lambda}\alpha - \left(1-\hat{\lambda}\right)\overline{\theta}\right) + \hat{\lambda}\beta x + \left(\hat{\lambda}\gamma + \left(1-\hat{\lambda}\right)/\gamma\right)\hat{p} + \left(\hat{\lambda}\gamma + \left(1-\hat{\lambda}\right)/\gamma\right)\hat{g} + \hat{f} \tag{28}$$

and

$$y = \left(\tilde{\lambda}\alpha - \left(1-\tilde{\lambda}\right)\overline{\theta}\right) + \tilde{\lambda}\beta x + \left(\tilde{\lambda}\gamma + \left(1-\tilde{\lambda}\right)/\gamma\right)\hat{p} + \left(\tilde{\lambda}\gamma + \left(1-\tilde{\lambda}\right)/\gamma + \tilde{\lambda}\phi\right)\hat{g} + \tilde{f} \tag{29}$$

However, since $\hat{g}$ is orthogonal to $\hat{p}$ and to x, its coefficient must be the same in (28) and (29). Hence, we have:

$$\tilde{\lambda}\gamma + \left(1-\tilde{\lambda}\right)/\gamma + \tilde{\lambda}\phi = \hat{\lambda}\gamma + \left(1-\hat{\lambda}\right)/\gamma \tag{30}$$

Rearranging the terms yields (27).

***Comment*:** Equation (27) provides a natural index of the contribution of the preference proxies used to elicit the policy maker's hidden information to improved identification. This is the product of the improvement in identification $\left(\tilde{\lambda} - \hat{\lambda}\right)$ between the OLS and the two-stage approaches times the maximum potential bias $\left(1-\gamma^2\right)/\gamma$ affecting the coefficients measuring the impact of p in (11) relative to (1). It may thus be called the Value Identification-Gap Narrowing Index (VIGNI), as it weighs the identification gap improvement by the upper bound of the potential bias at stake. Notice that the Hausman (1978) "exogeneity" test, as reformulated by Nakamura and Nakamura (1981), is precisely testing whether this coefficient is significant. Hence, it can also be interpreted as a deeper specification test that should systematically be used when preference proxies are used to narrow the identification gap of a policy trade off relative to OLS.

A closer look at (29) yields another useful result from a practical point of view, using in fact another argument from Nakamura and Nakamura (1981):

***Proposition* 5:** Applying OLS to (18) yields the same estimates as using 2SLS to estimate (7) using the preference proxies as instruments.

***Proof*:** Since $\hat{g}$ is orthogonal to $\hat{p}$ and to x and it has a zero mean, its inclusion in (29) does not affect the other estimates, which are in fact the 2SLS estimates as $\hat{p}$ is included instead of p in (29).

***Comments*:** The two-stage approach sketched above, which boils down equivalently to an application of the control-function approach or to 2SLS, provides the econometrician with good prospects of nearly identifying (1), and hence the impact of p on y. Moreover, this approach provides the natural test of improvement in identification by testing whether $\hat{g}$ is significant in (25). This is essentially what the Hausman "exogeneity" test does (Hausman, 1978; Nakamura & Nakamura, 1981). The key product of the first-stage equation is its estimated residuals series, which are orthogonal to the variables capturing the information that the policy maker and the econometrician have in common, in order to extract the signal of the unobserved information used by the former in making her decisions. This is what must guide the econometrician in his choice of instruments. The latter must be chosen as proxies for the policy maker's unobserved preference parameters. Hence, it is safer to use several instruments to make sure that the relevant unobserved information is captured without being contaminated by some trivial common information. This is likely to mitigate also the additional problem raised by the presence of ζ in $\hat{g}$, which makes the latter a noisy estimate of δe. This just entails a measurement error problem and hence a potential attenuation bias if $E(\zeta e) = 0$, without affecting identification as shown above. However, it is liable to bias the Hausman test toward zero, thus leading sometimes the econometrician to underestimate the true contribution of the preference proxies to the narrowing of the identification gap. Conversely, from an empirical point of view, this bias may reinforce the confidence that the econometrician may put in his near-identification strategy when the test turns out to be significant, despite the attenuation bias. Empirical examples of this approach are provided by Azam (2019) and Azam and Bhatia (2017) in a citizen's oversight perspective, using preference proxies that reveal intriguing aspects of some governments' preferences.

4. Conclusion

Unless the econometrician analyzes data produced in a randomized controlled trial, where it is known for sure that the policy has been applied at random, then the policy maker's behavior must be modeled to identify the policy trade off that the latter is supposedly exploiting. In this case, the model presented above has proved that using OLS will not in general yield the desirable identification, and thus might result in potentially highly misleading estimates and diagnosis. Fortunately, there is light at the end of the tunnel, as the econometrician can mitigate the asymmetric information problem by using some carefully selected instrumental variables. The latter must be useful proxies for the policy maker's underlying preference parameter(s). The first-stage equation is used to produce good residuals that reveal some of the policy maker's private information that is not correlated with the information that the econometrician shares with her. Because policy makers can have very weird preferences and will certainly not give away any accurate information about them in their speeches and writings, the econometrician might have to follow a drawn-out process of trial and errors. Fortunately, various forms of the Hausman exogeneity test can be used to evaluate whether the instruments used have made a significant contribution to the effort invested to come closer to identifying the policy trade off.

Funding

Funding from ANR under grant ANR-17-EURE-0010 (Investissement d'Avenir program) is gratefully acknowledged.

Acknowledgments

Helpful comments by Emmanuelle Auriol, Matteo Bobba, Sylvain Chabé-Ferret, Pierre Dubois and Claire Galez are gratefully acknowledged without implicating.

References

Angus. (2015). *Friends don't let Friends do IV*. Retrieved December 12, 2015, from http://mungowitzend.blogspot.com.es/2015/09/friends-dont-let-friends-do-IV.html?m=1

Arndt, C., Jones, S., & Tarp, F. (2015). Assessing Foreign Aid's Long Run Contribution to Growth and Development. *World Development*, *69*, 6-18. https://doi.org/10.1016/j.worlddev.2013.12.016

Azam, J.-P. (2019). Let's Call their Bluff: The Politics of Econometric Methodology. *Peace Economics, Peace Science, and Public Policy*, *25*(4), 1-7. https://doi.org/10.1515/peps-2019-0029

Azam, J.-P., & Bhatia, K. (2017). Provoking Insurgency in a Federal State: Theory and Application to India. *Public Choice*, *170*(3), 183-210. https://doi.org/10.1007/s11127-016-0389-5

Boone, P. (1996). Politics and the Effectiveness of Foreign Aid. *European Economic Review*, *40*, 289-329. https://doi.org/10.1016/0014-2921(95)00127-1

Desai, M. (1976). *Applied Econometrics*. Oxford: Philip Allan.

Easterly, W. (2006). *The White Man's Burden. Why the West's Effort to Aid the Rest Have Done so Much Ill and so Little Good*. New York: Penguin Press.

Fisher, F. M. (1965). Near Identifiability and the Variances of the Disturbance Terms. *Econometrica*, *33*, 409-419. https://doi.org/10.2307/1909798

Hausman, J. A. (1978). Specification Tests in Econometrics. *Econometrica*, *46*(6), 1251-1271. https://doi.org/10.2307/1913827

Moyo, D. (2009). *Dead Aid. Why Aid is not Working and How There Is Another Way for Africa*. London: Allen Lane.

Nakamura, A., & Nakamura, M. (1981). On the Relationships among Several Specification Tests Presented by Durbin, Wu and Hausman. *Econometrica*, *49*, 1583-1588. https://doi.org/10.2307/1911420

Stock, J. H., & Yogo, M. (2005). Testing for Weak Instruments in Linear IV Regression. In W. K. Donald, D. W. K. Andrews, & J. H. Stock (Eds.), *Identification and Inference for Econometric Models: Essays in Honor of Thomas Rothemberg* (pp. 80-108). New York: Cambridge University Press. https://doi.org/10.1017/CBO9780511614491.006

White, H., & Chalak, K. (2013). Identification and Identification Failure for Treatment Effects using Structural Systems. *Econometric Reviews*, *32*(3), 273-317. https://doi.org/10.1080/07474938.2012.690664

Note

Note 1. The classic reference on the near-identification issue is Fisher (1965) while Desai (1976) shows how to apply these concepts. White and Chalak (2013) address some of these issues in a much broader theoretical framework.